

使用多通道 EEG 识别和分类睡眠纺锤波

摘要

摘要:

本研究旨在复现一种高效准确的自动睡眠纺锤波检测方法。结合了基于滤波和阈值 (FBT) 的模型和机器学习模型。研究使用了公开的 DREAMS 数据集。

FBT 模型通过带通滤波、RMS 计算和多导联阈值检测来识别潜在的睡眠纺锤波。基于 FBT 模型的结果，复现了一个机器学习 pipeline，包括特征提取、自动模型选择和参数优化。实验结果表明，在 DREAMS 数据集上，本文选取的机器学习方法达到了接近人类专家的性能水平 (AUC 0.63 vs 0.73)。

该方法的优势在于结合了传统信号处理的效率和机器学习的适应性，能够有效处理多导联数据。然而，计算效率和对睡眠纺锤波亚型的区分仍有待进一步改进。

关键词：睡眠纺锤波、脑电图、自动检测、机器学习、信号处理

1 引言

睡眠纺锤波是在非快速眼动睡眠期间出现的特征性脑电图波形, 通常持续 0.5-2 秒, 频率在 11-16 Hz 之间。大量研究表明, 睡眠纺锤波与记忆巩固、认知功能以及睡眠质量密切相关。准确检测睡眠纺锤波对于理解睡眠的功能、评估睡眠质量以及诊断睡眠相关疾病具有重要意义。

然而, 由于睡眠纺锤波的特征变异性较大, 人工识别耗时费力且存在主观性, 自动检测睡眠纺锤波一直是脑电图 (EEG) 分析中的一个挑战性课题。传统的自动检测方法主要基于信号处理技术, 如带通滤波和阈值法, 但准确率有限。近年来, 随着机器学习技术的发展, 基于机器学习的睡眠纺锤波检测方法逐渐成为研究热点。

本研究旨在浮现一种高效准确的自动睡眠纺锤波检测方法。具体目标包括:

1. 构建一个基于滤波和阈值 (FBT) 的模型作为基线方法;
2. 基于 FBT 模型的结果, 开发和优化机器学习模型;
3. 比较 FBT 模型和机器学习模型的性能;
4. 在公开数据集上验证所提出方法的有效性。

通过这项研究, 希望能够复现一种快速、客观、准确的睡眠纺锤波检测工具, 帮助本人更快掌握相关领域的方法的使用流程, 培养本人的科研水平。

2 相关工作

睡眠纺锤波自动检测方法的研究已有数十年历史。早期的方法主要基于信号处理技术, 如 Schimicek 等人 [7] 提出的基于带通滤波和振幅阈值的方法。这类方法实现简单, 计算效率高, 但准确性有限。

随后, 研究者们尝试了更复杂的信号处理技术。例如 Huupponen 等 [4] 使用了自适应阈值来提高检测准确率。Tsanas 和 Clifford [8] 采用连续小波变换和局部加权平滑技术来检测睡眠纺锤波。这些改进在一定程度上提高了检测性能, 但仍存在局限性。近年来, 机器学习方法在睡眠纺锤波检测中显示出巨大潜力。Acir 和 Güzelis [1] 首次尝试使用支持向量机 (SVM) 来检测睡眠纺锤波。随后, 决策树 (Duman 等 [2])、人工神经网络 (Günes 等 [3]) 等方法也被应用到这一领域。这些方法通过学习大量标注数据的特征, 能够更好地适应睡眠纺锤波的变异性。

除了单一的机器学习模型, 一些研究者还尝试了混合方法。例如, Parekh 等 [6] 提出了一种基于稀疏优化的方法, 结合了信号处理和机器学习技术。O'Reilly 和 Nielsen [5] 比较了多种自动检测算法, 并强调了个性化方法的重要性。

尽管取得了显著进展, 睡眠纺锤波自动检测仍面临一些挑战:

1. 数据的个体差异性大, 难以建立通用模型;
2. 标注数据获取困难, 限制了监督学习方法的应用;
3. 多导联 EEG 数据的有效利用;
4. 实时检测的计算效率问题;
5. 与其他睡眠事件 (如 K 复合波) 的区分。

3 本文方法

3.1 本文方法概述

FBT 模型:

1. 使用 11-16 Hz 带通滤波和 RMS 计算检测纺锤波。
2. 基于阈值优化, 设置上下界参数, 通过网格搜索选择最优值。

机器学习管道:

1. 通过 FBT 模型生成的标签训练分类器。
2. 使用 Tree-based Pipeline Optimization Tool (TPOT) 自动选择最佳模型

本研究提出了一种两阶段的睡眠纺锤波检测方法, 包括基于滤波和阈值 (FBT) 的模型和基于机器学习的模型。

3.2 基于滤波和阈值 (FBT) 的模型

FBT 模型作为基线方法, 其主要步骤如下:

- a) 带通滤波: 对 EEG 信号进行 11-16Hz 的带通滤波, 这是睡眠纺锤波的典型频率范围。
- b) 计算 RMS: 使用高斯滑动窗口计算滤波后信号的均方根 (RMS) 值。

$$RMS[n] = \sqrt{\frac{\sum_{k=0}^n S[k] \cdot W[n-k]}{\text{sample size}}} \times C \quad (1)$$

其中 $S[k]$ 为信号, $W[n-k]$ 为高斯窗口核。

- c) 阈值检测: 计算 RMS 的修剪均值 () 和修剪标准差 (), 设定上下阈值:

$$\text{Lower boundary} = \mu_t(\text{RMS}) + T' \times \sigma_t(\text{RMS}) \quad (2)$$

$$\text{Higher boundary} = \mu_t(\text{RMS}) + T^h \times \sigma_t(\text{RMS}) \quad (3)$$

其中 T_l 和 T_h 分别为下阈值和上阈值参数。

- d) 多导联决策: 如果一个信号段在至少 3 个导联中满足阈值条件, 且持续时间在 0.5-2 秒之间, 则标记为潜在的睡眠纺锤波。

- e) 参数优化: 使用网格搜索和交叉验证优化 T_l 和 T_h 参数, 以最大化 ROC 曲线下面积 (AUC)。

3.3 基于机器学习的模型

基于 FBT 模型的结果, 设计了一个机器学习 pipeline:

- a) 数据分割: 使用 3 秒非重叠汉明窗口对 EEG 数据进行分割。
- b) 特征提取: 对每个分割提取三类特征:
 1. 局部阈值特征
 2. 峰值频率
 3. 峰值功率谱密度

- c) 特征标准化: 使用 Z-score 方法对特征进行标准化。
- d) 模型选择: 使用 Tree-based Pipeline Optimization Tool (TPOT) 自动选择最佳机器学习模型和超参数。
- e) 模型训练与评估: 使用 5 折交叉验证训练和评估模型, 以 AUC 为主要评估指标。

3.4 评价指标

使用以下指标比较 FBT 模型和机器学习模型的性能:

ROC 曲线和 AUC 值: ROC 曲线 (Receiver Operating Characteristic Curve) 通过展示分类器在不同阈值下的真阳率 (True Positive Rate, TPR) 和假阴率 (False Positive Rate, FPR) 的关系, 为评估模型整体分类性能提供了直观的方法。曲线下的面积 (AUC, Area Under Curve) 是衡量模型性能的标准化指标, AUC 值越接近 1.0, 表示分类器的综合性能越好。

混淆矩阵:

真阳性 (TP): 正确识别为纺锤波的样本数量。

假阳性 (FP): 被误分类为纺锤波的非纺锤波样本数量。

真阴性 (TN): 正确识别为非纺锤波的样本数量。

假阴性 (FN): 未能正确识别的纺锤波样本数量。

4 复现结果

4.1 与已有开源代码对比

我使用了原始代码中的主模型实现和数据处理流程, 并根据论文中描述的实验设置进行了修改以适应本地环境。

复现过程中的挑战与解决方案

1. 环境配置问题: 作者代码依赖较老的库版本, 我在复现中遇到了多个兼容性问题 (如 MNE 的版本老旧, 有些方法已经弃用)。通过逐步排查错误日志和参考社区建议, 我重新配置了兼容的运行环境。

2. 数据集处理: 原始数据集未提供明确的下载和处理方式。我根据论文细节重新实现了数据预处理流程, 确保数据能够正确输入模型。

5 实验结果分析

在 DREAMS 数据集上, 复现的方法与其他方法的比较结果如下 (平均 AUC \pm 标准差):

FBT 模型: 0.56 ± 0.085

机器学习模型: 0.63 ± 0.048

第二位专家: 0.73 ± 0.04

复现的机器学习模型性能接近人类专家, 且显著优于 FBT 模型。机器学习模型在各项指标上都优于 FBT 模型, 这表明学习数据特征可以提高检测准确性。选取的方法仍有一定的局限性: 计算效率仍有提升空间, 特别是对于长时间记录; 未考虑睡眠纺锤波的亚型 (如快纺锤波和慢纺锤波); 样本量相对有限, 可能影响机器学习模型的鲁棒性。总的来说, 这两个方法结合了传统信号处理和现代机器学习技术的优势, 在睡眠纺锤波检测任务上取得了良好的性能。

6 总结与展望

本次工作基于公开的源码，成功复现了论文中提出的睡眠纺锤波自动检测方法，并在实验中取得了接近论文结果的性能。使用基于带通滤波和阈值检测的方法预处理 EEG 信号，并提取潜在的纺锤波信号特征。构建了自动化的机器学习管道，包括特征提取、模型训练和参数调优，结合不同分类器（如支持向量机和随机森林）对纺锤波信号进行检测。在 DREAMS 数据集上验证了方法的性能，并取得了与人类专家接近的检测准确度（AUC 0.63 vs 0.73）。

尽管成功复现了论文方法，但在实现过程中仍存在以下不足之处：

1. 计算效率有待提升：复现的方法在处理长时间 EEG 记录时存在效率瓶颈。尤其是当信号时长显著增加时，带通滤波和 RMS 计算的耗时较长，限制了方法在大规模数据集上的实际应用。

2. 未考虑睡眠纺锤波的亚型：当前方法仅对纺锤波进行了整体检测，未区分快纺锤波和慢纺锤波两种亚型。然而，这些亚型可能对应不同的生理意义和功能，未来需进一步优化模型以实现更精细的分类。

3. 数据样本量有限：使用的数据集样本量相对较少，可能导致机器学习模型的鲁棒性不足。尤其是在测试集分布与训练集分布存在差异时，模型可能无法保持稳定的检测性能。

总的来说，本次工作结合了传统信号处理的高效性和现代机器学习的自适应性，在睡眠纺锤波检测任务上取得了良好的性能。然而，仍需在计算效率、亚型分类及数据扩展等方面进一步优化和研究，以推动睡眠纺锤波检测方法在实际场景中的广泛应用。

参考文献

- [1] Nurettin Acir and Cüneyt Güzeliş. Automatic spike detection in eeg by a two-stage procedure based on support vector machines. *Computers in Biology and Medicine*, 34(7):561–575, 2004.
- [2] Fazil Duman, Aykut Erdamar, Osman Erogul, Ziya Telatar, and Sinan Yetkin. Efficient sleep spindle detection algorithm with decision tree. *Expert Systems with Applications*, 36(6):9980–9985, 2009.
- [3] Salih Güneş, Mehmet Dursun, Kemal Polat, and Şebnem Yosunkaya. Sleep spindles recognition system based on time and frequency domain features. *Expert Systems with Applications*, 38(3):2455–2461, 2011.
- [4] E Huupponen, Alpo Värri, S Himanen, Joel Hasan, M Lehtokangas, and Jukka Saarinen. Optimization of sigma amplitude threshold in sleep spindle detection. *Journal of sleep research*, 9:327–34, 01 2001.
- [5] Christian O’Reilly and Tore Nielsen. Automatic sleep spindle detection: benchmarking with fine temporal resolution using open science tools. *Frontiers in Human Neuroscience*, 9, 2015.
- [6] Ankit Parekh, Ivan W. Selesnick, David M. Rapoport, and Indu Ayappa. Detection of k-complexes and sleep spindles (detoks) using sparse optimization. *Journal of Neuroscience Methods*, 251:37–46, 2015.
- [7] P. Schimicek, Prof. J. Zeitlhofer, P. Anderer, and B. Saletu. Automatic sleep-spindle detection procedure: Aspects of reliability and validity. *Clinical Electroencephalography*, 25(1):26–29, 1994.
- [8] Athanasios Tsanas and Gari D Clifford. Stage-independent, single lead eeg sleep spindle detection using the continuous wavelet transform and local weighted smoothing. *Frontiers in human neuroscience*, 9:181, 2015.